



Clustering algorithms with various types of tolerance

著者	?砂 幸裕
内容記述	Thesis (Ph. D. in Engineering)--University of Tsukuba, (A), no. 5318, 2010.3.25 Includes bibliographical references (leaves 88-94)
発行年	2010
URL	http://hdl.handle.net/2241/106237

氏 名 (本籍)	濱 砂 幸 裕 (宮 崎 県)		
学 位 の 種 類	博 士 (工 学)		
学 位 記 番 号	博 甲 第 5318 号		
学位授与年月日	平成 22 年 3 月 25 日		
学位授与の要件	学位規則第 4 条第 1 項該当		
審 査 研 究 科	システム情報工学研究科		
学 位 論 文 題 目	Clustering Algorithms with Various Types of Tolerance (様々な許容の概念によるクラスタリングアルゴリズム)		
主 査	筑波大学准教授	博士 (工学)	遠 藤 靖 典
副 査	筑波大学教授	工学博士	宮 本 定 明
副 査	筑波大学教授	工学博士	鬼 沢 武 久
副 査	筑波大学准教授	博士 (工学)	亀 山 啓 輔
副 査	筑波大学講師	博士 (工学)	延 原 肇

論 文 の 内 容 の 要 旨

データに伴う不確実性を扱うことはデータ解析の分野において盛んに議論されている重要なテーマである。一般に、誤差や欠損といった固有の不確実性を含んでいる場合、そのようなデータは空間上の点ではなく区間や集合として表されるが、通常は点で表されるデータを異なる形で扱うため、議論が複雑となるだけでなく、手法の構成も非常に難しくなる。そのような観点から、不確実性を伴うデータを従来と同じく空間上の点で扱う新たな手法として、許容範囲の概念が提案されている。

本論文では、データに伴う不確実性を扱う手法として、既に提案されている許容範囲の概念を用いた新たなクラスタリング手法の提案と、従来の許容にクラスタワイズという新たな概念を導入したクラスタワイズ許容の提案および、それを用いたクラスタリング手法の構築を行っている。はじめに、データに伴う不確実性を扱う手法として、従来の許容の概念を用いたクラスタリング手法を提案している。さらに、許容範囲付きデータに対する妥当性基準を提案している。妥当性基準とは、クラスタリング結果の評価や適切なクラス数数の決定に使われる指標であり、クラスタリング分野において重要なテーマの 1 つである。許容範囲の概念を用いた本手法では、各データを許容範囲付きデータとして扱い、不確実性を伴うデータを空間上の一点として扱うことが可能である。さらに、欠損値を含む心臓病の診断に関する実データを用いた数値例を示し、提案手法の有効性を示している。次に、クラスターの形状やサイズが異なるデータを扱う手法として、従来の許容の概念にクラスタワイズという新たな概念を導入したクラスタワイズ許容の概念を提案し、それを用いたクラスタリング手法を構築している。また、通常的手法では扱うことが困難なデータを用いた数値例を示し、提案手法の有効性を示している。

審 査 の 結 果 の 要 旨

本論文では、不確実性を伴うデータを扱う手法として、許容範囲の概念を用いたクラスタリング手法および妥当性基準を提案し、その有効性を示している。また、クラスターの形状やサイズの異なるデータを扱う手

法として、クラスタワイズ許容の概念を用いた手法を提案し、その有効性を示している。クラスタリングにおいて、データの持つ不確実性を最適化の枠組みで論じることのできる手法自体がほとんどないことに加え、その手法を発展させ、様々なクラスタ形状やサイズのデータを扱うことを可能とした本論文の手法は、新規性の観点から価値あるものとみなせる。また、本論文で展開されている最適化に関する数学的議論によって提案手法の信頼性は示されており、様々な数値例を通じてその有効性も検証されている。以上より、本論文で提案されている手法は、不確実性を伴うデータを扱うクラスタリングの新たな一手法として高く評価できる。今後の課題として、不確実性の意味・定式化に関して考察を深めるとともに、広く実データを通した検証が望まれる。

学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。